Ceph User Stories Highlights

Laura Flores, 2025/03/19

Introduction

Motivation

- Understand how Ceph is used in production
- Identify strengths and limitations
- Inform future improvements
- Share common knowledge among the user community as case studies

Survey overview

- Target audience: Ceph users running production clusters
- Total responses: **84** (**44** who shared contact info)
- Distributed via Slack, mailing list, and Meetup
- Open for ~ 3 weeks

Cluster Scale

How many production clusters are users managing?

- **60/84** manage 1-5 production clusters.
- **8/84** manage 6-10 production clusters.
- **16/84** manage 10+ production clusters.

What is the total size of these production clusters?

- The majority reported having greater than **1 PB** of total storage capacity. This was the largest choice offered in the survey.
- Some, using the “other” option, specified much larger numbers- up to **110 PB**.

How many OSDs and PGs per node are running on these production clusters?

- **6/84** reported **100+ total nodes** (Most reported a middle option, 11-50 total nodes).
- **36/84** reported **20+ OSDs per node** (this was the highest option on the survey).
- **3/84** reported **500+ PGs per OSD** (Most reported a middle option, 101–200 PGs per OSD).

Thoughts on a case study?

- There are some users who reported quite large clusters. We can focus case studies on these clusters and showcase their particular use cases and characteristics.

Deployment and Adding Storage

How are users deploying Ceph?

- **43/84** deploy on bare metal (RPMs/Binary)
- **35/84** deploy on containers (cephadm/Rook)
- **6/84** deploy on both

Have users added storage servers to their cluster?

- **75/84** said “yes”
- **9/84** said “no”

Do users find adding storage easy?

- Most users found adding storage easy, but some cited specific challenges such as time constraints, performance impacts, and preparation.

Thoughts on a case study?

- We can focus a case study on choosing a deployment method for a particular use case.
- Another could be about tips and challenges regarding storage addition in Ceph clusters.

Workloads and Applications

Which workloads are users running on their cluster?

- **72/84** are running a “cost & capacity optimized” workload
- **38/84** are running an “IOPS optimized” workload
- **34/84** are running a “throughput optimized” workload
- **19/84** are running a “latency sensitive” workload
- **1/84** specified that it depends on the type of cluster.

What solutions does Ceph provide for users?

- **61/84** use Ceph as a private cloud solution.
- **39/84** use Ceph for backup and recovery.
- **29/84** use Ceph for cold storage/archiving.
- … and more! But these were the most popular.

Thoughts on a case study?

- Many users wrote that they are looking to add a new use case in the future. It would be useful to have case studies about existing workloads and solutions to serve as guides to users who want to employ them in the future.

Monitoring and Benchmarking

What kinds of monitoring solutions are people using?

- Prometheus/Grafana (most common)
- Ceph Dashboard
- Zabbix
- CheckMK
- Nagios
- Icinga / Icinga2
- Centreon
- croit
- InfluxDB/Grafana
- TICK Stack
- Proxmox
- Custom scripts/internal systems
- … and more!

What benchmarking tools are people using to manage Ceph?

- FIO (most common)
- Warp
- radosbench
- CBT
- IO500
- Phoronix Test Suite
- vdbench
- sibench
- Locust
- Iperf3
- Custom tools
- … and more!

Thoughts on a case study?

- A case study could be made to highlight best practices with monitoring and benchmarking tools.

Upgrade Best Practices

How often are users upgrading their clusters?

- **38/84** upgrade every 3-12 months → this is more often than we thought!
- **27/84** upgrade every 1-2 years
- **8/84** upgrade as soon as a new release comes out
- **7/84** upgrade less than every two years
- **4/84** have never upgraded

What are the most popular upgrade methods?

- **45/84** upgrade with an orchestrator (cephadm/rook/ceph-ansible)
- **23/84** upgrade manually
- **12/84** use a custom/homemade upgrade method
- **4/84** have never upgraded

Thoughts on a case study?

- In addition to the above data, users wrote free responses on their upgrade best practices. There could be a case study focused on tips and tricks for various upgrade methods.

Feedback for Developers

What technological advancements have been most notable to users?

- Upgrading ease
- Rocksdb sharding
- Balancer
- PG scaling
- Cephadm
- Erasure coding
- Stretch clusters
- Bluestore improvements
- NVMe optimization
- Support of CephFS snapshots
- … and more!

What major feature are users expecting in the upcoming versions of Ceph?

- NFS port on Ceph Dashboard
- Crimson
- OSD Defragmenting
- NVMeOF
- Better handling of split-brain scenarios
- EC Improvements
- iscsi
- mclock fixes
- MSR
- Stability and overall bugfixes
- … and more!

Thoughts on a case study?

- Any case studies on the features above would be insightful, e.g. how people are using them, troubleshooting tips, suggested improvements, etc.

Next Steps

- The Ceph User Council will share survey results with Ceph Ambassadors in all regions.
- Ceph Ambassadors will work with local participants to create case studies.
- The participants who provided contact info (44/84) will be contacted for follow-up.
- Participants who did not share contact info are not eligible for case studies, but can contact the Ceph User Council if interested:
  - Laura Flores <lflores@ibm.com>
  - Neha Ojha <nojha@ibm.com>
  - Gaurav Sitlani <gsitlani@ibm.com>